

AD-A141 661

USE OF SIMULTANEOUS AUTOREGRESSIVE MODELS IN BALANCED  
INCOMPLETE BLOCK AN. (U) WISCONSIN UNIV-MADISON  
MATHEMATICS RESEARCH CENTER N R DRAPER ET AL. FEB 84

1/1

UNCLASSIFIED

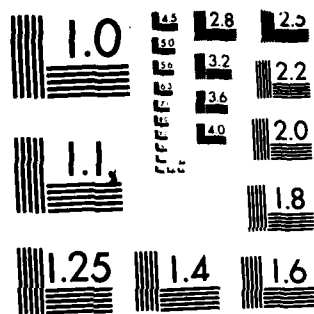
MRC-TSR-2650 DAAG29-80-C-0041

F/G 12/1

NL

|  |  |  |  |  |  |  |  |  |  |  |  |  |
|--|--|--|--|--|--|--|--|--|--|--|--|--|
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |
|  |  |  |  |  |  |  |  |  |  |  |  |  |

END  
DATE  
FILMED  
\*7-84  
DTIC



MICROCOPY RESOLUTION TEST CHART  
NATIONAL BUREAU OF STANDARDS-1963-A

AD-A141 661

MRC Technical Summary Report #2650

USE OF SIMULTANEOUS AUTOREGRESSIVE  
MODELS IN BALANCED INCOMPLETE BLOCK  
AND LATTICE SQUARE DESIGNS

N. R. Draper and D. Faraggi

Mathematics Research Center  
University of Wisconsin—Madison  
610 Walnut Street  
Madison, Wisconsin 53705

DTIC FILE COPY

February 1984

(Received January 27, 1984)

Sponsored by

U. S. Army Research Office  
P. O. Box 12211  
Research Triangle Park  
North Carolina 27709

Approved for public release  
Distribution unlimited

DTIC  
ELECTE  
MAY 31 1984

E

84 05 31 084

UNIVERSITY OF WISCONSIN-MADISON  
MATHEMATICS RESEARCH CENTER

USE OF SIMULTANEOUS AUTOREGRESSIVE MODELS IN BALANCED  
INCOMPLETE BLOCK AND LATTICE SQUARE DESIGNS

N. R. Draper and D. Faraggi

Technical Summary Report #2650  
February 1984

ABSTRACT

Simultaneous autoregressive models are applied to balance incomplete block and lattice square designs. Relationships between the resulting maximum likelihood estimators and standard analyses are explored and clarified. Also, it is shown that the Papadakis estimator, as applied to these designs, has the characteristics of the standard analyses, but simply involves fewer plots in the adjustment procedure for treatment averages. A worked example illustrates the differences in treatment estimates in a lattice square design, for standard inter-block analysis and for simultaneous autoregressive assumptions.

AMS (MOS) Subject Classification: 62K10

Key Words: Balanced incomplete block design, Lattice square design,  
Maximum likelihood estimators, Papadakis estimators,  
Simultaneous autoregressive models

Work Unit Number 4 (Statistics and Probability)

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041 and  
the Wisconsin Alumni Research Foundation via the UW Graduate School.

# SIGNIFICANCE AND EXPLANATION

The problem of analyzing data from field trials when observations in neighboring or nearby plots are correlated is an important one which has been extensively examined over the years. In 1937, Papadakis suggested a sensible but mysterious treatment estimator form whose role, and relationship to the corresponding maximum likelihood estimator has been debated for forty-odd years. These matters were clarified for certain one- and two-dimensional designs under certain simultaneous autoregressive models by Draper and Faraggi (1984). In the present paper, simultaneous autoregressive models are applied to balanced incomplete block and lattice square designs. Relationships between the resulting maximum likelihood estimators and standard analyses are explored and clarified. Also, it is shown that the Papadakis estimator, as applied to these designs, has the characteristics of the standard analyses, but simply involves fewer plots in the adjustment procedure for treatment averages. A worked example illustrates the differences in treatment estimates in a lattice square design, for standard inter-block analysis and for simultaneous autoregressive assumptions.

|                     |                                     |
|---------------------|-------------------------------------|
| Accession For       |                                     |
| NTIS GRA&I          | <input checked="" type="checkbox"/> |
| DTIC TAB            | <input type="checkbox"/>            |
| Unannounced         | <input type="checkbox"/>            |
| Justification       |                                     |
| By _____            |                                     |
| Distribution/ _____ |                                     |
| Availability Codes  |                                     |
| Dist                | Avail and/or<br>Special             |
| A-1                 |                                     |

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the authors of this report.

# USE OF SIMULTANEOUS AUTOREGRESSIVE MODELS IN BALANCED INCOMPLETE BLOCK AND LATTICE SQUARE DESIGNS

N. R. Draper and D. Faraggi

## 1. INTRODUCTION

The analysis of field trials when observations in nearby or neighbouring plots are correlated has received extensive study over a number of years. An early reference point is the work of Papadakis (1937). The relationship of his suggested estimator to the maximum likelihood estimators in certain one- and two-dimensional designs under certain simultaneous autoregressive models has been explored by Draper and Faraggi (1984).

It has been suggested by Kempton and Howes (1981, p. 65 and p. 69) that the Papadakis model may be used quite generally for field trials and specifically, for example, for lattice squares (LS's). On the other hand, a standard analysis for such designs under the "usual  $N(0, I\sigma^2)$ " type of error assumption is due to Yates (1936, 1939, 1940). Under Yates' analysis, adjustments are made to the treatment averages for the incompleteness of the blocks.

In this paper, we investigate lattice squares under the assumption that a simultaneous autoregressive model which assumes correlation within rows and columns is appropriate. This is done with different correlations,  $\rho_r$  for rows and  $\rho_c$  for columns. We then obtain maximum likelihood estimators for treatments and correlations and discuss the relationships and differences between the three possible treatments estimators namely maximum likelihood estimators, Yates', and Papadakis. It will be shown that, for these designs,

the Yates estimator is very close to the first order term in an expansion, where valid, of the maximum likelihood estimator. Moreover, the Yates' and Papadakis corrections to the treatment means have similar forms. However, neither is actually needed in designs completely balanced for rows and columns because an explicit solution can be obtained. An example is presented for illustration in Section 5.

The lattice square is a two-dimensional design. As a preliminary, we investigate what can be regarded as a one-dimensional equivalent, namely the balanced incomplete block design. Results and conclusions which closely parallel those of the lattice square are obtained.

## 2. BALANCED INCOMPLETE BLOCK DESIGNS

Suppose we wish to examine  $t$  treatments in  $b$  blocks of size  $k < t$ . Suppose, further, that each treatment occurs the same number,  $m$ , of times and appears  $\lambda$  times with every other treatment. Then the design is said to be a balanced incomplete block design (BIBD). Note that there are  $n = bk = tm$  plots, and that  $\lambda(t-1) = m(k-1)$ . For the standard details of analysis reproduced below see, for example, Cochran and Cox (1957), or Yates (1940). The usual model considered is

$$y_{js} = \mu + \alpha_s^* + \beta_j + \epsilon_{js} \quad (2.1)$$

if "replicates" are not a factor; if they are, some minor differences occur in some formulas. Here  $y_{js}$  is the observation on treatment  $s$  in block  $j$ ,  $\mu$  is an overall mean,  $\alpha_s^*$  is the effect of the  $s$ th treatment,  $\beta_j$  is the effect of the  $j$ th block and  $\epsilon_{js}$  is an error such that  $\epsilon = (\text{column vector of } \epsilon_{js} \text{ in a defined order}) \sim N(0, I\sigma^2)$ . Two alternative assumptions on the  $\beta_j$  are commonly made.

(a)  $\beta_j$  is a fixed effect. This leads to the so-called intra-block analysis.

(b)  $\beta = (\beta_1, \dots, \beta_b)' \sim N(0, I\sigma_\beta^2)$  independently of the  $\epsilon_{js}$ . This leads to the analysis with the recovery of inter-block information. As  $\sigma_\beta^2 \rightarrow \infty$ , case (b)  $\rightarrow$  case (a).

We set



$B_s$  = total of all response observations  $y$  from  
blocks in which treatment  $s$  appears.

$T_s$  = total of all observations on treatment  $s$ .

$G = \sum_s T_s$  = grand total ( $= \sum_s B_s$ ).

Then, from Cochran and Cox (1957, pp. 445-446) we can estimate the effect of the  $s$ th treatment by

$$\hat{\alpha}_s^* = m^{-1} \{T_s + \hat{\theta}[(t-k)T_s - (t-1)B_s + (k-1)G]\} \quad (2.2)$$

where  $\hat{\theta}$  is an adjustment factor taking the value

$$\hat{\theta} = \{t(k-1)\}^{-1} \quad (2.3)$$

in case (a) when inter-block information is not recovered, and the value

$$\hat{\theta} = (b-1)(E_b - E_e) / \{t(k-1)(b-1)E_b + (t-k)(b-t)E_e\} \quad (2.4)$$

where it is recovered. The quantities  $E_b$  and  $E_e$  are the mean squares due to blocks adjusted for treatments, and residual, respectively. Define the  $t$  by  $bk = n$  matrix  $T$  as follows. When the blocked design is listed in order, block by block,  $T_{si} = 1$  if the  $s$ th treatment is associated with the  $i$ th observation, and  $T_{si} = 0$  otherwise. Let

$$D_k = 11' - I_k \quad (2.5)$$

where  $1' = (1, 1, \dots, 1)$  of dimension 1 by k. Then define

$$D = I_b \otimes D_k \quad (2.6)$$

where  $\otimes$  denotes the Kronecker product. Then, in matrix form, with  $\alpha^* = (\alpha_1^*, \dots, \alpha_s^*, \dots, \alpha_t^*)'$ , (2.2) becomes

$$\hat{\alpha}^* = m^{-1} \{TY - \hat{\theta}(t-1)TD[I - m^{-1}T'T]Y\} \quad (2.7)$$

where  $Y$  is the vector of observations, recorded in the same pattern as the block by block enumeration.

We now consider, instead of case (a) or case (b) above, an alternative model  $y_i = \alpha_s + x_i$ , where

$$x_i = \rho \sum_{j \neq i} n_{ij} x_j + \epsilon_i \quad (2.8)$$

and where  $n_{ij} = 1$  if  $y_i$  and  $y_j$  occur in the same block, and  $n_{ij} = 0$  otherwise. We assume  $\epsilon = (\epsilon_1, \dots, \epsilon_n)' \sim N(0, I\sigma^2)$  independently of the  $x$ 's. Following Besag's (1974, Eq. (4.13) with  $\mu = 0$ ) simultaneous autoregressive model, we obtain the likelihood

$$(2\pi\sigma^2)^{-n/2} |B| \exp\{-(2\sigma^2)^{-1} X'B'BX\}, \quad (2.9)$$

where  $B = I - \rho D$  and  $X = Y - T'\alpha$ . The maximum likelihood estimator  $\hat{\alpha}$  is of form (Ripley, 1981, (5.34))

$$(TV^{-1}T')^{-1}TV^{-1}Y, \quad (2.10)$$

where  $TV^{-1}T' = T(I - \hat{\rho}D)^2T' = m\{I - m^{-1}T(2\hat{\rho}D - \hat{\rho}^2D^2)T'\}$  using the fact that  $TT' = mI$ . The matrix  $(TV^{-1}T')^{-1}$  can be expanded as a convergent series if and only if the eigenvalues of  $Q_B = m^{-1}T(2\hat{\rho}D - \hat{\rho}^2D^2)T'$  are all less than one in absolute value. This is true if

$$(1 - 2^{1/2})/(k-1) < \hat{\rho} < \begin{cases} \min\{(1+2^{1/2}), t^{1/2}-1\}, & k = 2, \\ \min\{\frac{1}{2}(1+2^{1/2}), \frac{1}{4}(t-1)\}, & k = 3, \\ (1+2^{1/2})/(k-1), & k \geq 4. \end{cases} \quad (2.11)$$

If we retain only terms of orders zero and one in  $Q_B$ , we obtain

$$\hat{\alpha}_1 = m^{-1}[TY - \{2\hat{\rho} - \hat{\rho}^2(k-2)\}TD(I - m^{-1}T'T)Y] \quad (2.12)$$

where we have used the facts that  $D^2 = (k-1)I_n + (k-2)D$ , and  $T(I - m^{-1}T'T) = 0$ . Comparing (2.12) with (2.7) we notice the estimators are identical except that  $\hat{\theta}(t-1)$  in (2.7) replaces  $\{2\hat{\rho} - \hat{\rho}^2(k-2)\}$  in (2.12). Because of the validity of the expansion,  $\hat{\alpha}_1$  will converge to the maximum likelihood estimator  $\hat{\alpha}$ , for fixed  $\hat{\rho}$ , if the former is applied iteratively in the following sense:

$$\hat{\alpha}_{j+1} = m^{-1}[TY - \{2\hat{\rho} - \hat{\rho}^2(k-2)\}TD(Y - T'\hat{\alpha}_j)] \quad (2.13)$$

where  $\hat{\alpha}_0 = m^{-1}TY$ . In fact, iteration is not needed here because the balanced incomplete block design produces a  $TV^{-1}T'$  of patterned form which can be

explicitly inverted (Rao, 1973, p. 67) to give  $\hat{\alpha}$  in terms of  $\hat{\rho}$  as

$$\hat{\alpha} = a(I + bTDT')T(I - \hat{\phi}D)Y \quad (2.14)$$

where  $a = \{m - (t-2)\hat{\phi}\} / [\{m - (t-1)\hat{\phi}\}\{m + \hat{\phi}\}]$ ,  $b = \hat{\phi} / \{m - (t-2)\hat{\phi}\}$  and  $\hat{\phi} = \{2\hat{\rho} - (k-2)\hat{\rho}^2\} / \{1 + (k-1)\hat{\rho}^2\}$ . The maximum likelihood  $\hat{\rho}$  is obtained as follows. Setting the first derivative of the log likelihood with respect to  $\rho$  equal to zero provides

$$\frac{(Y - T'\hat{\alpha})'(D - \hat{\rho}D^2)(Y - T'\hat{\alpha})}{(Y - T'\hat{\alpha})(I - 2\hat{\rho}D + \hat{\rho}^2D^2)(Y - T'\hat{\alpha})} = \frac{(k-1)\hat{\rho}}{\{1 + (n-1)\hat{\rho}\}\{1 + \hat{\rho}\}} \quad (2.15)$$

The  $\hat{\alpha}$  may be substituted from (2.14) to provide an equation in  $\hat{\rho}$  which can be solved, for example, via the Newton-Raphson method. From this,  $\hat{\alpha}$  is then obtained. In practice, it is easier to replace  $\alpha$  in the likelihood by  $\hat{\alpha}$  from (2.14) and then to maximize the likelihood numerically by a search over  $\hat{\rho}$ .

For least squares estimation, the term on the right of (2.15) is replaced by zero, so that

$$\hat{\rho}_{LS} = (Y - T\hat{\alpha}_{LS})'D(Y - T\hat{\alpha}_{LS}) / \{(Y - T\hat{\alpha}_{LS})'D^2(Y - T\hat{\alpha}_{LS})\}, \quad (2.16)$$

where  $\hat{\alpha}_{LS}$  has the same form as  $\hat{\alpha}$  but with  $\hat{\rho}$  replaced by  $\hat{\rho}_{LS}$ . Again numerical maximization is simplest.

### 3. LATTICE SQUARES

Suppose we wish to examine  $t$  treatments in a set of  $m$   $k$  by  $k$  lattice squares. Typically  $m = \delta(k+1)$  where  $\delta = \frac{1}{2}, 1, 2, \dots$ . The different values of  $\delta$  correspond to various association relationships among the treatments. When  $\delta = \frac{1}{2}$ , any treatment  $s$  will appear with any other treatment  $w$  once in either a row or a column but not both. When  $\delta = 1$ ,  $s$  and  $w$  appear together both in a row and a column, once each, and so on. The usual model considered is

$$y_i = \mu + \alpha_s^* + \pi_q + \beta_r + \gamma_c + \epsilon_i \quad (3.1)$$

where the set of subscripts  $(q,r,c)$  representing the position in the  $r$ th row and the  $c$ th column within the  $q$ th replicate have been replaced by a single subscript  $i = 1, 2, \dots, n$  which denotes the plot's position in a row by row enumeration of the plots, replicate by replicate. Two alternative assumptions on the  $\beta$ 's and the  $\gamma$ 's are commonly made

(a)  $\beta_r, \gamma_c$  are fixed effects. This leads to an intra-block analysis.

(b)  $\beta_r \sim N(0, \sigma_\beta^2)$  and  $\gamma_c \sim N(0, \sigma_\gamma^2)$  and all  $\beta$ 's and  $\gamma$ 's are independent.

Typically it is assumed that  $\sigma_\gamma^2 = \sigma_\beta^2$ , although this is not necessary. As  $\sigma_\beta^2 \rightarrow \infty$  and  $\sigma_\gamma^2 \rightarrow \infty$ , case (b)  $\rightarrow$  case (a).

We set, for  $\delta \geq 1$ ,

$T_s$  = total of all response observations  $y$  from all plots  $i$   
which receive treatment  $s$ ,

$R_s$  = total of all  $y$ 's in rows in which treatment  $s$  appears,

$C_s$  = total of all  $y$ 's in columns in which treatment  $s$  appears,

$G$  = grand total of all  $y$ 's,

$L_s = (m-1)T_s - mR_s + G$ ,

$M_s = (m-1)T_s - mC_s + G$ .

Then, from Cochran and Cox (1957, p. 493), we can estimate the effect of the  $s$ th treatment by

$$\hat{\alpha}_s^* = m^{-1} \{T_s + \lambda_r L_s + \lambda_c M_s\}, \quad (3.2)$$

where  $\lambda_r$  and  $\lambda_c$  are coefficients for the adjustments needed for the incompleteness of the rows and columns. The values of  $\lambda_r$  and  $\lambda_c$  depend on the replication selected and which case, (a) or (b) is assumed. Examples are in Appendix A.

For  $\delta = \frac{1}{2}$ , the formula for  $\hat{\alpha}_s^*$  is of the same form as (3.2) but with redefined symbols. We do not discuss this because our matrix formulation does not contain this ambiguity.

Define the  $t$  by  $n$  matrix  $T$  as follows. When the blocked design is written out row by row within replicate and then replicate by replicate,  $T_{si} = 1$  if the  $s$ th treatment is associated with the  $i$ th observation, and  $T_{si} = 0$  otherwise. Let  $D_k = 11' - I_k$ , where  $1' = (1, 1, \dots, 1)$  of dimension

1 by k, and define

$$\begin{aligned} Z_r &= I_m \otimes I_k \otimes D_k, \\ Z_c &= I_m \otimes D_k \otimes I_k, \\ Z &= Z_r + Z_c, \end{aligned} \quad (3.3)$$

where  $\otimes$  denotes the Kronecker product. Then, in matrix form, with  $\alpha^L = (\alpha_1^L, \alpha_2^L, \dots, \alpha_t^L)'$ , the L denoting lattice, (3.2) becomes

$$\hat{\alpha}^L = m^{-1} \{TY - mT(\lambda_r Z_r + \lambda_c Z_c)(I - m^{-1}T'T)Y\} \quad (3.4)$$

where Y is the vector of observations, recorded in the same pattern as described earlier.

We now consider, instead of case (a) or case (b), an alternative model  $y_i = \alpha_s + x_i$ ,  $i = 1, \dots, n$ ,  $s = 1, \dots, t$ , where

$$x_i = \rho_r \sum_{j \neq i} n_{ij}^{(r)} x_j + \rho_c \sum_{j \neq i} n_{ij}^{(c)} x_j \quad (3.5)$$

and where  $n_{ij}^{(r)} = 1$  if  $y_i$  and  $y_j$  occur in the same row and replicate and is zero otherwise, while  $n_{ij}^{(c)} = 1$  if  $y_i$  and  $y_j$  occur in the same column and replicate and is zero otherwise. We assume  $\epsilon \sim (\epsilon_1, \dots, \epsilon_n)' \sim N(0, I\sigma^2)$  independently of the x's. Following Besag's (1974, Eq. (4.13) with  $\mu=0$ ) simultaneous autoregressive model, we obtain the likelihood  $(2\pi\sigma^2)^{-n/2} |B| \exp\{-(2\sigma^2)^{-1} X'B'BX\}$  where  $B = I - \rho_r Z_r - \rho_c Z_c$ .

$X = Y - T'\alpha$ . The maximum likelihood estimator  $\hat{\alpha}$  is of form  $(TV^{-1}T')^{-1}TV^{-1}Y$ , where  $TV^{-1}T' = T(I - \hat{\rho}_r Z_r - \hat{\rho}_c Z_c)^2 T' = m(I - Q_L)$  and where

$$Q_L = m^{-1}T(2\hat{\rho}_r Z_r + 2\hat{\rho}_c Z_c - \hat{\rho}_r^2 Z_r^2 - \hat{\rho}_c^2 Z_c^2 - 2\hat{\rho}_r \hat{\rho}_c Z_r Z_c)T'. \quad (3.6)$$

The matrix  $(TV^{-1}T')^{-1}$  can be expanded as a convergent series if and only if the eigenvalues of  $Q_L$  are all less than one in absolute value. This is true if

$$\begin{aligned} -1 < q_1 + (k^2 - 1)q_2 < 1, \\ -1 < q_1 - q_2 < 1, \end{aligned}$$

where

$$\begin{aligned} q_1 &= -(m-2)(\hat{\rho}_r^2 + \hat{\rho}_c^2), \\ q_2 &= \{2\hat{\rho}_r + 2\hat{\rho}_c - (m-3)(\hat{\rho}_r^2 + \hat{\rho}_c^2) - 2(k-1)\hat{\rho}_r \hat{\rho}_c\} / (k+1). \end{aligned} \quad (3.7)$$

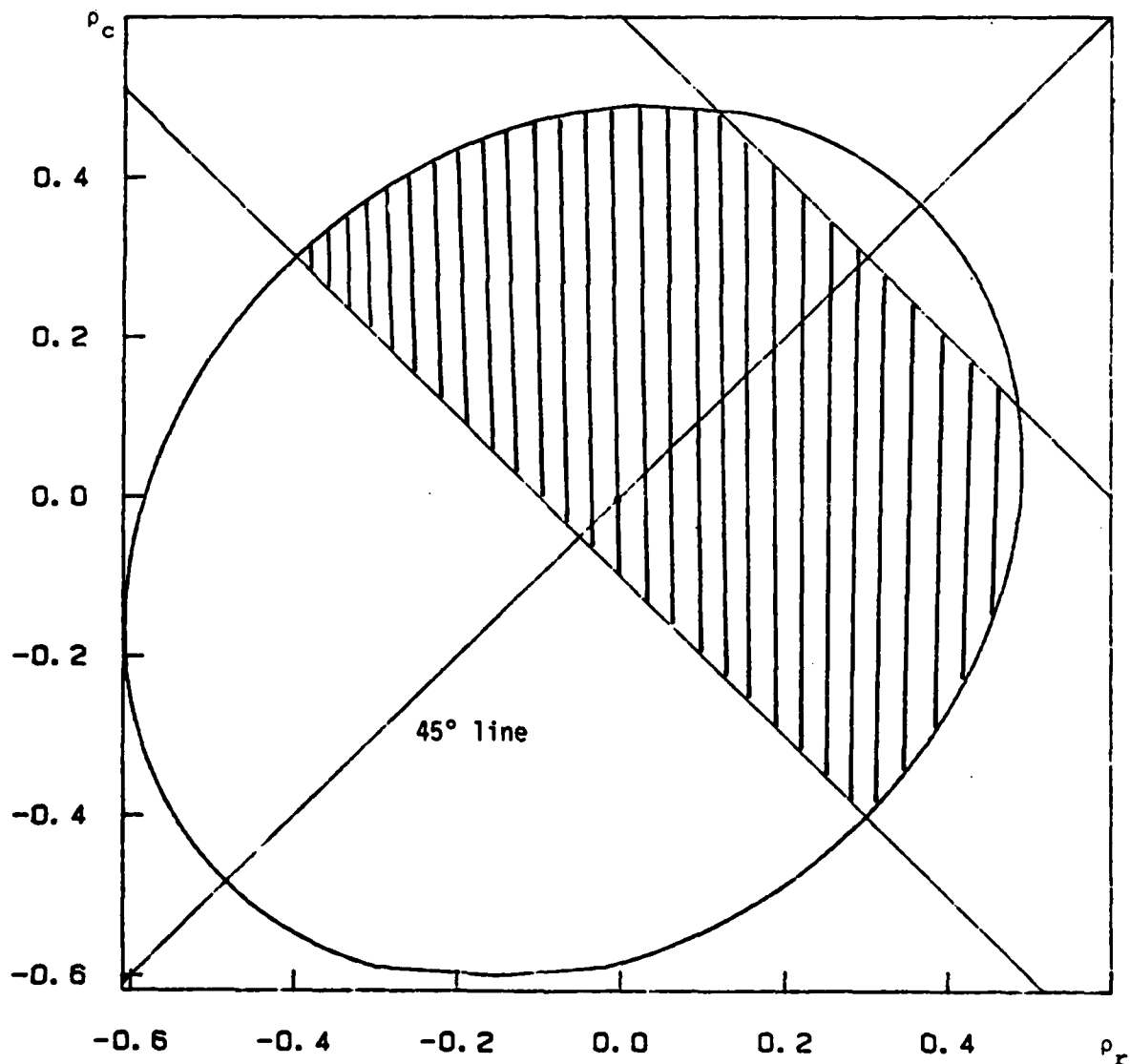
For example, if  $m = 6$ ,  $k = 5$ , the conditions imply a feasible region shown shaded in Figure 1.

If we retain only terms of orders zero and one in  $Q_L$ , we obtain

$$\hat{\alpha}_1^L = m^{-1}[TY - T(\hat{\phi}_r Z_r + \hat{\phi}_c Z_c)(I - m^{-1}T'T)Y], \quad (3.8)$$



Figure 1. The shaded region shows the values of  $(\rho_r, \rho_c)$  for which the expansion of the maximum likelihood estimator involving  $Q_L$  is valid, for the case  $m = 6, k = 5$  described in the text. For the case  $\rho_r = \rho_c = \rho$  this region reduces to the intersection of the shaded region and the  $45^\circ$  line shown, with a multiplicative scaling factor of  $2^{-1/2}$  applied.



where  $\hat{\phi}_u = 2\hat{\rho}_u - (m-3)\hat{\rho}_u^2 + 2\hat{\rho}_r\hat{\rho}_c$  and  $u = r, c$  in turn. In deriving  $\hat{\alpha}_1^L$  we have used the following facts ( $u=r$  or  $c$ ):

$$\begin{aligned} Z_u^2 &= (m-2)I + (m-3)Z_u, \\ TZ_r Z_c &= 1_t 1_t' T - TZ_r - TZ_c \\ T(I - m^{-1}T'T) &= 0, \\ 1_n'(I - m^{-1}T'T) &= 0. \end{aligned} \tag{3.9}$$

Comparing (3.8) with (3.4), we see that the estimators are identical except that  $m\lambda_u$  in (3.4) replaces  $\hat{\phi}_u$  in (3.8). Because of the validity of the expansion,  $\hat{\alpha}_1^L$  will converge to the maximum likelihood estimator  $\hat{\alpha}^L$ , for fixed  $\hat{\rho}_r$  and  $\hat{\rho}_c$ , if  $\hat{\alpha}_1^L$  is applied iteratively in the following sense

$$\hat{\alpha}_{j+1}^L = m^{-1}\{TY - T(\hat{\phi}_r Z_r + \hat{\phi}_c Z_c)(Y - T'\hat{\alpha}_j^L)\}, \tag{3.10}$$

where  $\hat{\alpha}_0^L = m^{-1}TY$ . In fact, iteration is not needed here for  $\delta \geq 1$ , because the lattice design produces a  $TV^{-1}T'$  of patterned form which can be explicitly inverted (Rao, 1973, p. 67) to give  $\hat{\alpha}^L$  in terms of  $\hat{\rho}_r$  and  $\hat{\rho}_c$  as

$$\{(a-b)I + b11'\}T(I-Q)Y, \tag{3.11}$$

where

$$\begin{aligned} a &= \{(m-1)q_1 - (k^2-2)mq_2\} / \{[(m-1)q_1 - (k^2-1)mq_2]\{(m-1)q_1 + mq_2\}\}, \\ b &= mq_2 / \{[(m-1)q_1 - (k^2-1)mq_2]\{(m-1)q_1 + mq_2\}\}. \end{aligned} \quad (3.12)$$

The maximum likelihood  $\hat{\rho}_r$  and  $\hat{\rho}_c$  are obtained as follow. Setting the first derivatives of the log likelihoods with respect to  $\rho_1$  and  $\rho_2$  equal to zero provides equations:

$$\frac{n\hat{X}'Z_u\hat{B}\hat{X}}{\hat{X}'\hat{B}^2\hat{X}} + \frac{\left\{ \frac{\partial}{\partial \rho_u} |B| \right\}_{\rho_r=\hat{\rho}_r, \rho_c=\hat{\rho}_c}}{|\hat{B}|} = 0, \quad u = r, c \quad (3.13)$$

where  $\hat{X}$  and  $\hat{B}$  are the values of  $X$  and  $B$  with  $\hat{\alpha}^L$ ,  $\hat{\rho}_r$  and  $\hat{\rho}_c$  replacing  $\alpha$ ,  $\rho_r$  and  $\rho_c$ . These two simultaneous equations involve, in their second terms, the  $(i, i+1)$  and the  $(i, i+k)$  elements, the designations being reduced modulo  $n$  where they exceed  $n$ , respectively of  $\hat{B}^{-1}$  and must be solved, a tedious calculation. It is simpler in practice to substitute a selected grid of  $(\hat{\rho}_r, \hat{\rho}_c)$  values into  $\hat{\alpha}^L$  and hence evaluate the likelihood over that grid, so picking out the maximum value of the likelihood.

For least squares estimation, the term on the right of (3.13) is replaced by zero so that

$$\begin{bmatrix} \hat{\rho}_{r,LS} \\ \hat{\rho}_{c,LS} \end{bmatrix} = \begin{bmatrix} b_{rr} & b_{rc} \\ b_{cr} & b_{cc} \end{bmatrix}^{-1} \begin{bmatrix} X'Z_rX \\ X'Z_cX \end{bmatrix} \quad (3.14)$$

where  $b_{uv} = X'Z_uZ_vX$ . The least squares estimate of  $\alpha$ ,  $\hat{\alpha}_{LS}^L$  is of the same form as  $\hat{\alpha}^L$  but has  $\hat{\rho}_u$  replaced by  $\hat{\rho}_{u,LS}$  everywhere.

Special Case  $\rho_r = \rho_c = \rho$ .

In many practical situations, it is reasonable to assume that  $\rho_r = \rho_c = \rho$ . For comments on this choice, see Kempton and Howes (1981, p. 63). All formulas reduce to their appropriate forms via simple substitution of  $\rho_r = \rho_c = \rho$ . All that is needed is to write  $Z = Z_r + Z_c$  whenever it occurs and to drop subscripts from  $\lambda$ ,  $\phi$ , and  $\rho$ . The equation parallel to Eqs. (3.13) is the result of adding these two equations together.

#### 4. LATTICE SQUARE DESIGN EXAMPLE

We shall re-examine data given by Cochran and Cox (1957, pp. 490-493) comprising a lattice square design with  $k = 4$ ,  $\delta = 1$ ,  $m = 5$ ,  $n = 80$ . The appropriate formulas for Yates' estimators  $\lambda_r$  and  $\lambda_c$  are given in Appendix A and here result in the numerical values  $\lambda_r = 0.04787$  and  $\lambda_c = 0.03037$  (p. 493). The maximum likelihood estimates for  $\rho_r$  and  $\rho_c$  are  $\hat{\rho}_r = 0.108$  and  $\hat{\rho}_c = -0.030$  and these lead to estimates for the 16 treatment parameters given in the second column of Table 1. These may be compared with the "adjusted means" of Cochran and Cox (1957, p. 491) which are the corresponding Yates' estimates, shown as the third column of Table 1. We see that, although two entirely different models are used, individual estimates are broadly speaking, compatible. In view of the fact that only one at most of these models can be correct for the data set, we do not compare them via a model-dependent criterion such as weighted or unweighted sum of squares of residuals.

Table 1. Comparison of treatment effects for a 4 by 4 lattice square design with five replications. Second column: Maximum likelihood estimates under the assumption of a simultaneous autoregressive model. Third column: Yates' estimates from Cochran and Cox (1957, p. 491).

| Treatment | Mle     | Yates'  |
|-----------|---------|---------|
| 1         | 4.98    | 6.45    |
| 2         | 12.71   | 13.68   |
| 3         | 8.90    | 8.73    |
| 4         | 11.41   | 11.36   |
| 5         | 9.89    | 9.44    |
| 6         | 5.96    | 7.58    |
| 7         | 7.52    | 7.37    |
| 8         | 9.79    | 9.32    |
| 9         | 10.91   | 10.01   |
| 10        | 15.77   | 14.91   |
| 11        | 18.12   | 17.59   |
| 12        | 12.91   | 12.70   |
| 13        | 12.07   | 10.69   |
| 14        | 13.40   | 14.27   |
| 15        | 10.07   | 9.28    |
| 16        | 10.05   | 11.09   |
| Sum       | 174.46* | 174.47* |

\*These figures are subject to small rounding errors; both should be 174.48.

## 5. PAPADAKIS AND MAXIMUM LIKELIHOOD

We reconsider the assumption made in connection with Eq. (3.1).

Suppose that now  $\beta_r$  is the combined neighbouring effect from the "plot to the left in a row", and the "plot to the right in a row", these left and right adjacencies being determined in torus, wrap-around manner (Martin, 1982). A similar "up and down in a column" definition is applied now to  $\gamma_c$ . Let  $N_r$  and  $N_c$  be the  $n$  by  $n$  neighbour-specification matrices for rows and columns whose  $i$ th row contains 1 in positions  $j$  for which plot  $j$  is row- and column-adjacent respectively, to plot  $i$ , and zero otherwise. Then the maximum likelihood estimator of  $\alpha^*$  is

$$\hat{\alpha}_p^* = m^{-1} \{TY - mT(\psi_r N_r + \psi_c N_c)(I - m^{-1}T'T)Y\} \quad (5.1)$$

where  $\psi_r$  and  $\psi_c$  are coefficients for the adjustments needed for the row- and column-adjacency effects. Parallel to the work in Section 3, there will be two forms for  $\psi_r$  and  $\psi_c$  according to whether an intra-neighbour analysis ( $\beta_r, \gamma_c$  fixed effects) or an inter-neighbour analysis ( $\beta_r, \gamma_c$  random effects) is specified.

If we compare (5.1) with Eq. (3) of Draper and Faraggi (1984), we see that they would be equivalent if  $m\psi_r = m\psi_c = \hat{\phi}$  and if  $N = N_r + N_c$ . In fact, (5.1) is the Papadakis estimator for the case  $\psi_r \neq \psi_c$ , up to definition of  $\psi_r$  and  $\psi_c$ . The Papadakis formulation would involve regression estimators of  $\psi_r$  and  $\psi_c$  instead. We thus see that the Papadakis estimator can be regarded, up to definition of  $\psi_r$  and  $\psi_c$ , as a maximum likelihood estimator in the particular model formulation described above.

## 6. REMARKS

A previous paper (Draper and Faraggi, 1984) discussed field trials in one and two dimensions for designs of Type II(a), Type III, and Type III cyclically row permuted (Williams, 1952) for one and two dimensional layouts. In the two dimensional layout, no directional differences were assumed and a single  $\rho$  value was used for both rows and columns. The extension to  $\rho_r$  and  $\rho_c$  for that case is easily effected using formulas of the type exhibited in this paper. The maximum likelihood estimator has the form of (2.10), with

$$V^{-1} = (I - \rho_r N_r - \rho_c N_c)^2 \quad (6.1)$$

where  $N_r$  is an  $n$  by  $n$  row-neighbour-specification matrix defined by

$$N_r = I_{\delta(k+1)} \otimes I_k \otimes N, \quad (6.2)$$

where  $N_c$  is an  $n$  by  $n$  column-neighbour-specification matrix defined by

$$N_c = I_{\delta(k+1)} \otimes N \otimes I_k, \quad (6.3)$$

and where  $N$  is the cyclic  $k$  by  $k$  matrix

$$N = \begin{bmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 1 \\ 1 & 0 & 1 & 0 & \dots & 0 & 0 \\ \dots & & & & & & \\ 1 & 0 & 0 & 0 & & & 1 & 0 \end{bmatrix}. \quad (6.4)$$



We can re-write (2.10) as

$$m^{-1}(I - m^{-1}TQT')^{-1}T(I-Q)Y \quad (6.5)$$

where

$$Q = 2\hat{\rho}_r N_r + 2\hat{\rho}_c N_c - \hat{\rho}_r^2 N_r^2 - \hat{\rho}_c^2 N_c^2 - 2\hat{\rho}_r \hat{\rho}_c N_r N_c. \quad (6.6)$$

Here,

$$N_r^2 = 2I + N_r^{(2)}, \quad N_c^2 = 2I + N_c^{(2)}, \quad N_r N_c = N_{rc}^{(11)},$$

where  $N_r^{(2)}$ ,  $N_c^{(2)}$ , and  $N_{rc}^{(11)}$  denote  $n$  by  $n$  neighbour-specification matrices which contain 1's and 0's such that

$$N_r^{(2)} = I_{\delta(k+1)} \otimes I_k \otimes N^{(2)} \quad (6.7)$$

$$N_c^{(2)} = I_{\delta(k+1)} \otimes N^{(2)} \otimes I_k \quad (6.8)$$

$$N_{rc}^{(11)} = I_{\delta(k+1)} \otimes N \otimes N \quad (6.9)$$

where  $N^{(2)}$  is the cyclic  $k \times k$  matrix

$$N^{(2)} = \begin{bmatrix} 0 & 0 & 1 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & \dots & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & \dots & 0 & 0 & 0 \\ \dots & & & & & & & & \\ 0 & 1 & 0 & 0 & \dots & 1 & 0 & 0 \end{bmatrix} \quad (6.10)$$

and  $N$  is defined in (6.4). For example, the  $i$ th row of  $N_r^{(2)}$  has 1's in those columns  $j$  for which cell  $j$  is located next but one to cell  $i$  in a row of the design, and has zeros otherwise. Table 2 illustrates the general situation and has those (row) cells marked with symbol  $N_r^{(2)}$ . Similar remarks apply for  $N_c^{(2)}$  and  $N_{rc}^{(11)}$  for columns and diagonals respectively, as shown in Table 2. When cell  $i$  is near or on an edge, the appropriate torus-generated cells must be used.

|             |                 |             |                 |             |
|-------------|-----------------|-------------|-----------------|-------------|
|             |                 | $N_c^{(2)}$ |                 |             |
|             | $N_{rc}^{(11)}$ | $N_c$       | $N_{rc}^{(11)}$ |             |
| $N_r^{(2)}$ | $N_r$           | Cell $i$    | $N_r$           | $N_r^{(2)}$ |
|             | $N_{rc}^{(11)}$ | $N_c$       | $N_{rc}^{(11)}$ |             |
|             |                 | $N_c^{(2)}$ |                 |             |

Table 2. Cell patterns which lead to the  $n$  by  $n$  neighbour-specification matrices  $N_r$ ,  $N_c$ ,  $N_r^{(2)}$ ,  $N_c^{(2)}$ ,  $N_{rc}^{(11)}$ .

The expansion of (6.5), when valid, as far as terms of zero and first order in  $Q$ , can be written

$$\begin{aligned} & m^{-1}\{TY - TQ(I - m^{-1}T'T)Y\} \\ &= m^{-1}\{TY - T(2\hat{\rho}_r N_r + 2\hat{\rho}_c N_c - \hat{\rho}_r^2 N_r^{(2)} - \hat{\rho}_c^2 N_c^{(2)} - 2\hat{\rho}_r \hat{\rho}_c N_{rc}^{(11)})(I - m^{-1}T'T)Y\}, \quad (6.11) \end{aligned}$$

where we have used the fact that  $T(I - m^{-1}T'T) = 0$ .

## REFERENCES

- BESAG, J. (1974). Spatial interaction and the statistical analysis of lattice systems (with discussion). J. R. Statist. Soc. B 36, 192-236.
- COCHRAN, W. G. and COX, G. M. (1957). Experimental Designs, 2nd edition. New York: Wiley.
- DRAPER, N. R. and FARAGGI, D. (1984). Neighbouring plot analysis in one and two dimensions. Biometrika 00, 000-000.
- KEMPTON, R. A. and HOWES, C. W. (1981). The use of neighbouring plot values in the analysis of variety trials. Appl. Statist. 30, 59-70.
- MARTIN, R. J. (1982). Some aspects of experimental design and analysis when errors are correlated. Biometrika 69, 597-612.
- PAPADAKIS, J. S. (1937). Méthode statistique pour des expériences sur champ. Bull. Inst. Amél. Plantes à Salonique, No. 23.
- RIPLEY, B. D. (1981). Spatial Statistics. New York: Wiley.
- RAO, C. R. (1973). Linear Statistical Inference and Its Applications, 2nd edition. New York: Wiley.
- WILLIAMS, R. M. (1952). Experimental designs for serially correlated observations. Biometrika 39, 151-167.
- YATES, F. (1936). Incomplete randomised blocks. Ann. Eugenics London 7, 121-140.
- YATES, F. (1939). The recovery of inter-block information in variety trials arranged in three-dimensional lattices. Ann. Eugenics 9, 136-156.
- YATES, F. (1940). The recovery of inter-block information in balanced incomplete block designs. Ann. Eugenics 10, 317-325.

Appendix A. Values of  $\lambda_r, \lambda_c$  for  $\delta = \frac{1}{2}, 1, 2$  for  
Lattice Square Designs.

| $\delta$      | <u>With recovery of<br/>inter-block information</u>        |  | <u>Without recovery of<br/>inter-block information</u> |
|---------------|--|--|--|
|               | $\lambda_r$  | $\lambda_c$  | $\lambda_r = \lambda_c$                                |
| $\frac{1}{2}$ | $\frac{2(E_r - E_e)}{k(k+1)E_r}$                           | $\frac{2(E_c - E_e)}{k(k+1)E_c}$                           | $\frac{2}{k(k+1)}$                                     |
| 1             | $\frac{(E_r - E_e)(kE_c - E_e)}{(k-1)(k^2E_rE_c - E_e^2)}$ | $\frac{(E_c - E_e)(kE_r - E_e)}{(k-1)(k^2E_rE_c - E_e^2)}$ | $\frac{1}{(k-1)k}$                                     |
| 2             | $\frac{W_e - W_r}{k(W_r + W_c + (k-1)W_e)}$                | $\frac{W_e - W_c}{k(W_r + W_c + (k-1)W_e)}$                | $\frac{1}{(k-1)k}$                                     |

In the above table,

$$W_e = E_e^{-1}, W_r = (2k-1)/(2kE_r - E_e), W_c = (2k-1)/(2kE_c - E_e),$$

$E_r$  = mean square due to rows adjusted for treatments,

$E_c$  = mean square due to columns adjusted for treatments and rows,

$E_e$  = residual mean square.

| REPORT DOCUMENTATION PAGE  |                                     | READ INSTRUCTIONS<br>BEFORE COMPLETING FORM  |
|--|-------------------------------------|--|
| 1. REPORT NUMBER<br>2650   | 2. GOVT ACCESSION NO.<br>AD-A141661 | 3. RECIPIENT'S CATALOG NUMBER  |
| 4. TITLE (and Subtitle)<br><br>USE OF SIMULTANEOUS AUTOREGRESSIVE MODELS IN<br>BALANCED INCOMPLETE BLOCK AND LATTICE SQUARE<br>DESIGNS   |                                     | 5. TYPE OF REPORT & PERIOD COVERED<br>Summary Report - no specific<br>reporting period                               |
|  |                                     | 6. PERFORMING ORG. REPORT NUMBER   |
| 7. AUTHOR(s)<br><br>N. R. Draper and D. Faraggi  |                                     | 8. CONTRACT OR GRANT NUMBER(s)<br><br>DAAG29-80-C-0041   |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Mathematics Research Center, University of<br>610 Walnut Street Wisconsin<br>Madison, Wisconsin 53706   |                                     | 10. PROGRAM ELEMENT, PROJECT, TASK<br>AREA & WORK UNIT NUMBERS<br>Work Unit Number 4 -<br>Statistics and Probability |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>U. S. Army Research Office<br>P. O. Box 12211<br>Research Triangle Park, North Carolina 27709   |                                     | 12. REPORT DATE<br>February 1984   |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)  |                                     | 13. NUMBER OF PAGES<br>24  |
|  |                                     | 15. SECURITY CLASS. (of this report)<br><br>UNCLASSIFIED   |
|  |                                     | 15a. DECLASSIFICATION/DOWNGRADING<br>SCHEDULE  |
| 16. DISTRIBUTION STATEMENT (of this Report)<br><br>Approved for public release; distribution unlimited.  |                                     |  |
| 17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)   |                                     |  |
| 18. SUPPLEMENTARY NOTES  |                                     |  |
| 19. KEY WORDS (Continue on reverse side if necessary and identify by block number)<br><br>Balanced incomplete block design, Lattice square design, Maximum likelihood<br>estimators, Papadakis estimators, Simultaneous autoregressive models  |                                     |  |
| 20. ABSTRACT (Continue on reverse side if necessary and identify by block number)<br><br>Simultaneous autoregressive models are applied to balance incomplete<br>block and lattice square designs. Relationships between the resulting maximum<br>likelihood estimators and standard analyses are explored and clarified. Also,<br>it is shown that the Papadakis estimator, as applied to these designs, has the<br>characteristics of the standard analyses, but simply involves fewer plots in<br>the adjustment procedure for treatment averages. A worked example illustrates<br>the differences in treatment estimates in a lattice square design, for<br>standard inter-block analysis and for simultaneous autoregressive assumptions. |                                     |  |

**DAT  
FILM**